White Paper Report

Report ID: 103982

Application Number: HD5137211

Project Director: William Cowan (wgcowan@indiana.edu)

Institution: Indiana University, Bloomington

Reporting Period: 9/1/2011-8/31/2013

Report Due: 11/30/2013

Date Submitted: 2/25/2014

White Paper

HD-51372-11

Incorporating Annotated Video into Omeka

William George Cowan

Indiana University, Bloomington

21 February 2014

White Paper

Using Scholarly Video

Most sites on the web that support digital video use what might be termed the entertainment model for playing back digital video. Basically the entertainment model is that the user wants to start at the beginning of the video and play the video to the end. The user may want to stop the video and restart, even scrub to other parts of the video and usually wants to resume the video at a given point if the user leaves the site and comes back.

Some sites may offer the ability to bookmark a specific place in a video but they don't generally provide the ability to have extensive annotation with the bookmark. In fact I've noticed many times when videos are embedded in a site, the commentary often includes a phrase like "check out what happens at 4:50" to make sure the viewer sees the important point in the video.

What I am attempting to accomplish with using video in Omeka is to give the scholar, archivist, staff member or owner of the video, the ability to segment the video with a specific start time and end time for the segment, to be able to annotate that segment, to be able to either playback that segment by itself or to play it back within the context of the longer video. Most users that I have talked with would like to be able to do this without resorting to Adobe Premier or Final Cut Pro to have to actually cut the video into segments and work with those precut segments but would rather have the video streamed and to be able to designate start and end points in the streaming video. In addition to annotation, users want to add transcription and translation, be able to tag a segment and be able to search and discover similar segments that can then be played back.

When this project was originally proposed, I thought the best way to accomplish the playback of these video segments was by designing themes to support that playback. And I still believe that for some uses of video this is the best choice. But many organizations don't have the staff or technical support to build themes but must use existing themes. Consequently, our ideas for playing video in Omeka evolved over the course of the project and we developed the VideoStream plugin which uses an existing theme and plugs the video playback element automatically into the show item page.

Project Activities and Accomplishments

During the project period, the project went through several different phases as we discovered how to effectively use streaming video in Omeka.

There are several assumptions we are making in the design for using video in Omeka. First, that many scholarly uses of video involve small segments of video, not the presentation of an entire video file. That it is beneficial to learning, discovery and research that these video segments have additional annotation associated with them. And finally, that most researchers, archivists and curators working with Omeka do not want to use a video editing tool, like Adobe Premier, to segment their video and have to create individual Omeka items from each of these distinct video segments.

Consequently our design uses the Annotator's Workbench, an external tool for digital video segmentation and annotation, to provide an XML file that can be used to load video segments as items into Omeka.

Another issue for digital video is how to access the video file. We had originally looked at loading the video file into the Omeka file system and just doing a progressive download of

the video. This approach has some disadvantages mainly in that it limits the setting of the start and end points in viewing the video segment. By using a video streaming server instead and streaming the video to the web site, you can set start and end points in the video and work more interactively with the player using javascript.

The design for loading segmented, annotated video into Omeka involved working in three areas.

The first was to ensure that Omeka would have the necessary fields to support video segments. Currently Omeka's Dublin Core fields do not support start and stop points in a video file in order to access

The second was the design and development of a plugin to support loading video segments and annotations from the Annotator's Workbench (AWB) XML file. There is already a plugin, CsvImport, that will load comma, separated values from a file into Omeka. Each row in the file becomes an item in Omeka. To handle the AWB files, which are just XML files, we modified the CsvImport plugin to use a XSLT stylesheet to transform the AWB XML file into a comma, separated value file that could then use the rest of the CsvImport code to import this CSV file into Omeka without having to develop any new code. So the first result of this project is the XmlImport plugin which will take any XML file, let the user choose an XSLT file that will transform the XML to a csv file and then proceed to load that csv file as items into Omeka. While this process works for the AWB XML file, it can work for any XML that has a proper XSLT to generate the csv file.

The third area was creating an Omeka theme that would support the display of the annotated video segments within Omeka. The theme developed provides code in the items directory of the theme that displays video segments properly as well as other items that are not video segments. In addition, the theme properly displays the individual video segments when you click on the link that goes to an individual object. This was done by modifications to the browse.php and show.php in the items directory in the theme. By doing the modifications here, this overrides the default browse and show functions and uses the customized code.

So the first phase involved the original goal of the project, the development of a plugin that would allow the breakdown of the XML file, created by the Annotator's Workbench, into individual items that would represent the annotated video segments that were created using this tool. By making a modification to the existing Comma Separated Value plugin that already existed for Omeka, we were able to create a plugin that would take the XML file from Annotator's Workbench, apply a XSLT style sheet to that XML file to create a CSV file that could be loaded into Omeka. To make this work, we needed to make manual modifications to the Moving Image Data Type to add fields for the video filename, streaming server name, start point in the video and the end point in the video. After making these changes to Omeka and installing the XML plugin, we were able to load a standard Annotator's Workbench file, transform with an XSLT file to create the CSV file and then load those items into Omeka as annotated video segments, with start and stop points defined in the Moving Image type. So we now had items in Omeka that represented annotated segments of a streaming video. However, Omeka at that time was not able to easily playback the video so we needed to change Omeka to handle the play back of these video segments.

The next phase of the project focused on web-publishing, how to effectively playback the video in Omeka, especially the video segments we have created with the XML from Annotator's Workbench. The first pass we made at this was to modify the themes in Omeka. Consequently, we created a theme in Omeka that would look for items that had a video segment start value in the field in Moving Images data type and if that field was not null, then display a small video playback window next to the item in the result list. In addition, when

you clicked on the item in the result list, on the show item page, we displayed a larger video playback screen.

The major flaw with this approach has to do with the current state of video players for the web. While nearly every video player allows you to start the video at a particular time code, there is no restriction, using the player's video controls, to scrubbing back in the video to a point earlier in the larger streaming video. And not all video players handle an end point the same way. Some pay not attention to it at all and simply keep playing past the end point. Others end the video at that point and don't show any additional video beyond that end point. Since the start point is not restricted to showing as starting at the start point, neither of the implementations for the end point work, It becomes very unclear which segment is being played and what exactly the start and stop points are.

As work on the design of these pages continued, it became apparent that for scholarly purposes, the playback of video that has been segmented and annotated is more complex than most video playback on the web.

For instance, does the scholar want the user to have access to the entire content of the video so that the user can scrub back and forth in the video to areas not originally part of the segment? Or should the user be restricted to just the segment that is revealed to them and must use other means than scrubbing through the video to go to other segments and annotations? Depending on the collection, either of these methods might be the best. We have designed themes in Omeka that support either of these scenarios.

In addition, decisions have to be made about how the user is able to navigate to different segments and annotations. By incorporating other plugins for Omeka, like the Simile Timeline plugin, we have successfully set up interactive timelines that display the various, available segments and allow the user to navigate to any of those segments.

To illustrate these various design features, we have implemented a web site at http://www.dlib.indiana.edu/projects/omekavideo that illustrate these different issues with video display as well as provide access to the plugins and themes that we have developed over the last year or so to support video segmentation and annotation in Omeka.

This work with themes has been beneficial in other ways. I began collaboration with Richard Edwards, and to a lesser degree with Shannon Clute, based on a book on Film Noir called *The Maltese Touch of Evil: Film Noir and Potential Criticism* (Dartmouth College Press/UPNE, 2011), which they co-authored. In this book, they discussed the need for a database for what they termed noiremes, small segments of the film that represent what the film is about and what film noir is about. Working with Richard Edwards, we put together a database using principles of the *OULIPO* literary movement in the hope generating a nearly limitless analysis based on each film being divided into 100 segments based on the overall length of the film. For instance, if a film was 100 minutes long, each segment would be 1 minute.

Working with 20 public domain films noir from the Internet Archive and using the techniques from the development of themes for video segmentation, we were able to create an Omeka site that shows all 100 segments for a given film, with a player to playback that segment. In addition, using the ability of Omeka to tag items, each video segment has a tag representing the particular segment. So segment 33 in each film is tagged as 33 so that you can say for these 20 films, show me the 33rd segment or the 33% point in the movie. Making these side by side comparisons allows a completely different kind of analysis of film and being able to not only see a photo of that segment, but also being able to playback that segment for multiple films is also unique. See http://www.dlib.indiana.edu/projects/oufinopo for an example of how this technique works.

At this time, we also started to consider the needs of smaller organizations or institutions that either did not have the technical resources to design or re-design Omeka themes to support video or the funds to hire staff or consultants to do the work. In addition, the Center for History and New Media at George Mason University released a major upgrade to Omeka, going from 1.5.3 to 2.0. This major upgrade required the rewriting of existing code for the project and a redefinition of some of the goals of the project.

So we begin the next phase by looking at designing and implementing a new video streaming plugin that can display annotated and segmented digital video in Omeka without requiring the creation or modification of any themes. In addition, I didn't want to have the manual modification of the Moving Image Data type that working with themes in the previous phase required.

The video streaming plugin we began designing and implementing at this time would automatically have a video playback component that would be automatically loaded by the default theme and the two basic themes that were provided with Release 2.0 of Omeka. Also at the time the plugin is installed, it automatically creates a Streaming Video data element set that contains the fields needed for streaming annotated and segmented video that displays with the item on the administrative side of the Omeka site. And the plugin is easily configurable to display back the video in a player that can be sized, will play back only a specific segments if desired and if the necessary video files are transcoded, will play back on Firefox, IE, Safari or Chrome as well as on Apple iPad tablets and iPhones. At this time, there are still issues with playing back video on Android devices due to changing versions of the operating system and the types of files supported on Android tablets and phones. See the Appendix for screen shots of the interface for the Video Streaming Plugin.

The new plugin we have developed has also allowed us to be more flexible with the source of the streaming video as well. For instance, we have tested the plugin using the Avalon Media System under development here at IU and YouTube, as well as an Adobe Media Server for flash and HLS formatted delivery for iOS devices. In addition I have streamed video and annotated and segmented it for use on an Omeka site for IU's internal video podcast system.

The plugin is becoming known in the Omeka community and last fall we did a major social media release of information about the plugin and how to access it. In addition, I have gone to 5 different THAT Camps, one was related to the Oral History Association annual meeting and another was related to the Digital Library Federation annual meeting, and made presentations on the work done under this grant. At the Oral History Association Annual meeting in Cleveland, I had a booth in the vendor's display area for two days and showed and discussed the use of the plugins we had developed. I have done 6 presentations at the university on using this plugin and using Omeka in general over the course of the grant period.

Collaborations

While we accomplished the goals outlined in our proposal, the development of an Omeka plugin to load the annotated video segments created in the Annotator's Workbench as Items in Omeka, our work in this area showed that to make the project successful we would need to go beyond just developing this plugin. So in addition to loading segments from AWB, we also developed techniques for displaying video segments and annotation in Omeka first by modifying or creating themes and then by creating a Video Streaming plugin that will work with any theme.

In addition our work with other audiences than the original ethnographers whose field work videos were part of the Ethnographic Video for Instruction and Analysis Digital Archive that developed and used the Annotator's Workbench showed how flexible and powerful these

simple video techniques could be. For instance, with Richard Edwards and his work with Film Noir, we were able to create a complex display of video with an underlying structure supporting his concept of a Noireme database in less than a week. In another instance, working with a T'ai Chi instructor, in several days, I was able to synchronize still images of a training session with a YouTube video of a T'ai Chi master to show the different positions.

As mentioned earlier, the original audience was the users of the Annotator's Workbench, which was part of the EVIADA project. But as we developed different techniques for displaying the video, from themes to plugins, we also discovered different audiences, from film studies to instructors in non-traditional settings to archivists.

In particular I am currently working with three different groups on three different implementations of Omeka that are using the Video Streaming Plugin to display video segments and annotations. The first group is the English Department at Indiana University. Don Belton, an English professor at IU Bloomington, recently passed away. To honor him, the department wanted a web site to show his books and other artifacts of his work at the university. They used Omeka since it could support the variety of media they wanted to show, specifically some videos of his talks and lectures. This site is at http://belton.indiana.edu/ and under collections shows some videos of Don Belton that use the Video Streaming Plugin that we developed through this grant to display the video.

Second, Brian Graney, the archivist for the Black Film Center / Archive at Indiana University, Bloomington was able to secure an NEH Grant for a conference and workshop in November 2013. At the workshop, the various sessions and presentations were recorded and those recordings, along with other proceedings of the conference, will be available on line through Omeka and the Video Streaming plugin we developed with this grant. The Omeka site with the proceedings is not yet available but the site about the conference is at http://www.indiana.edu/~regener8/regener8/.

Third, I am working with Lynn M. Hooker, Associate Professor of Hungarian Studies at IU, who has received grant monies to give a symposium on a "Virtual Romani Cultural House" this summer. This symposium will also make use of the work we have done on Omeka and video.

Potential Uses

The successful use of Omeka on the Indiana University campus and at other institutions and organizations around the country, coupled with the continued increase in the use and importance of video both for instruction and research make the use of tools like the Video Streaming plugin important to the Omeka Community.

In the past several months I have been thinking more about the pedagogical uses of Omeka and in particular this video plugin. I have already discussed some, perhaps, more obvious uses of the work we have done on this grant with ethnographic field videos, film studies and lecture or conference playback with annotation. I have also thought this would be beneficial in language learning, being able to playback the video of a given conversation, with transcription, translation and annotation of the nuances of the language, would seem to present possibilities in self learning. In clinical psychology education one could tape sessions, annotate them, play back for student learning. In addition with the segmentation and annotation, search and discovery would allow the students to review similar segments that discussed similar problems with different patients allowing a comparison of symptoms and diagnosis. For music, students could play back different performances of the same score, perhaps, side by side to discover nuances in the performances that are not readily visible in separate playback. In STEM education there are probably hundreds, if not thousands of

videos on YouTube that could be enhanced considerably by annotation. This plugin would allow such use of YouTube.

Currently, the VideoStream plugin sets certain configuration options at the site level for the size and location of the video playback. An option for the future that I am working on is to have those options set at the item level so that for a given item you would be able to override the site level options for size, playback, etc.

Available Products from this Project

There are three main products from this grant. First is a plugin for Omeka that allows the ingestion of the XML file created by Annotator's Workbench into Omeka with the video segments and annotations from AWB becoming items in Omeka. This is accomplished by using an XSLT stylesheet to transform the AWB XML file into a comma separated value file that is then loaded using the CSV code. This plugin only works with Omeka 1.5. This plugin is available at http://www.dlib.indiana.edu/projects/omekavideo. However, this general technique, using an XSLT to transform an XML file, could be applied externally to Omeka, transform an XML file into a CSV file; then the CSV file created in the previous step could be loaded into Omeka through the CsvImport plugin. Because of this we decided not to create an XmlImport plugin for Omeka 2.0 or greater.

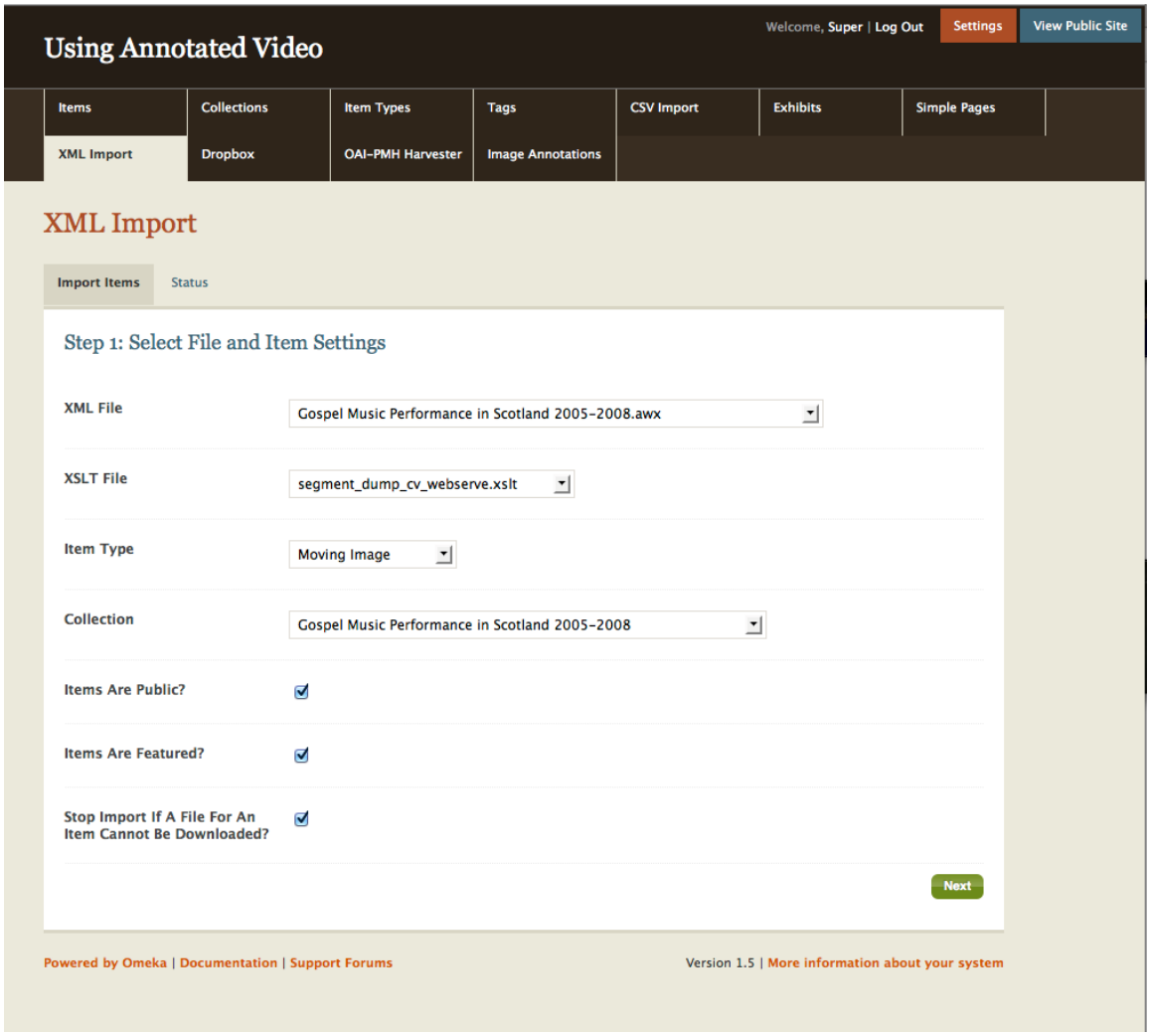
The next product was a set of themes that can represent items that are segmented and annotated streaming video. The website http://www.dlib.indiana.edu/projects/omekavideo shows how theses different themes can be used to present streaming video segments in Omeka. In addition, the website for film noir, http://www.dlib.indiana.edu/projects/oufinopo also shows how themes can be used to present even more complex presentations of video segments in Omeka.

Finally, the VideoStream plugin allows the display of the video, based on configuration options in the plugin, anywhere on an Omeka page by using the *fire_plugin_hook* for *public_items_show* on that page. This allows video playback for any theme with little or no changes to the theme. This plugin is designed primarily to work with Omeka 2.0 or greater, but a more limited version of the plugin is available for Omeka 1.5. This plugin is available at the http://www.dlib.indiana.edu/projects/omeka2 site.

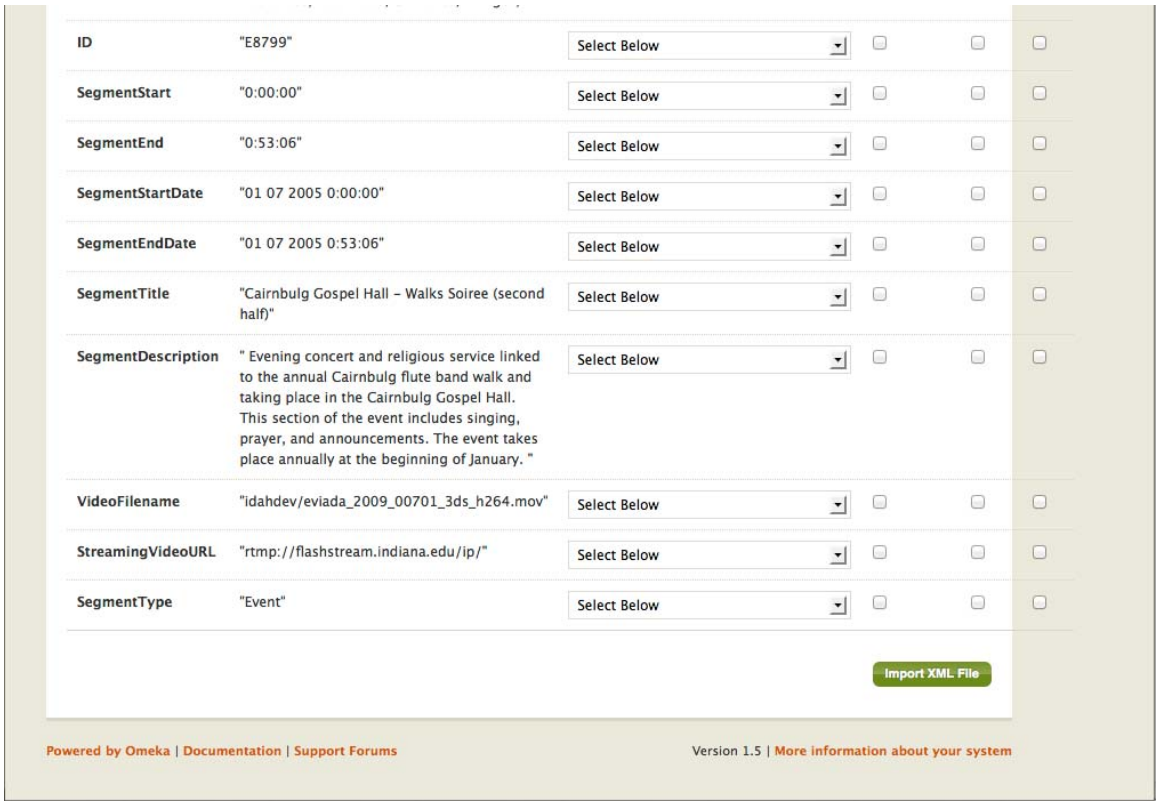
Documentation is available on all the sites referenced above for using the various tools mentioned in this paper.

Appendix

The appendix shows a series of web pages from Omeka with the various products from this grant.



First page for XmlImport plugin. Select the XML file and the transforming XSLT file.



Second page of XmlImport plugin for mapping fields in the CSV file to fields in Dublin Core in Omeka.

Different themes for viewing segments. This theme allows you to playback the entire video but displays information about the segment you are in.

Title
The Gospel Male Voice Choir sing 'The King and I'

Subject
English language, Electronic piano, Keyboard instruments, Listening, Oral communication, Vocal music, Sacred vocal music, Choral music, Gospel music, Choral conducting, Choral societies, Sacred songs, Conducting, Music, Laudatory songs, Hymns, Songs, Gospel singers, Singers, Choral societies, Vocal groups, Musical groups, Vocal ensembles, Males, Men, Baptists , Applause, Audiences, Gesture,

Description
The North of Scotland GMVC stand and sing 'The King and I', which is taken from Mosie Lister's Rise up and Sing.

Date
01 07 2005 0:23:52

Identifier
S1512

Coverage
01 07 2005 0:27:22

Items in Collection
Reality sing 'Alpha and Omega' (beginning missing) [0:00:20--0:04:47]

The North of Scotland GMVC perform two songs [0:04:47--0:16:20]

New Life sing 'Sweet Peace' [0:16:20--0:21:11]

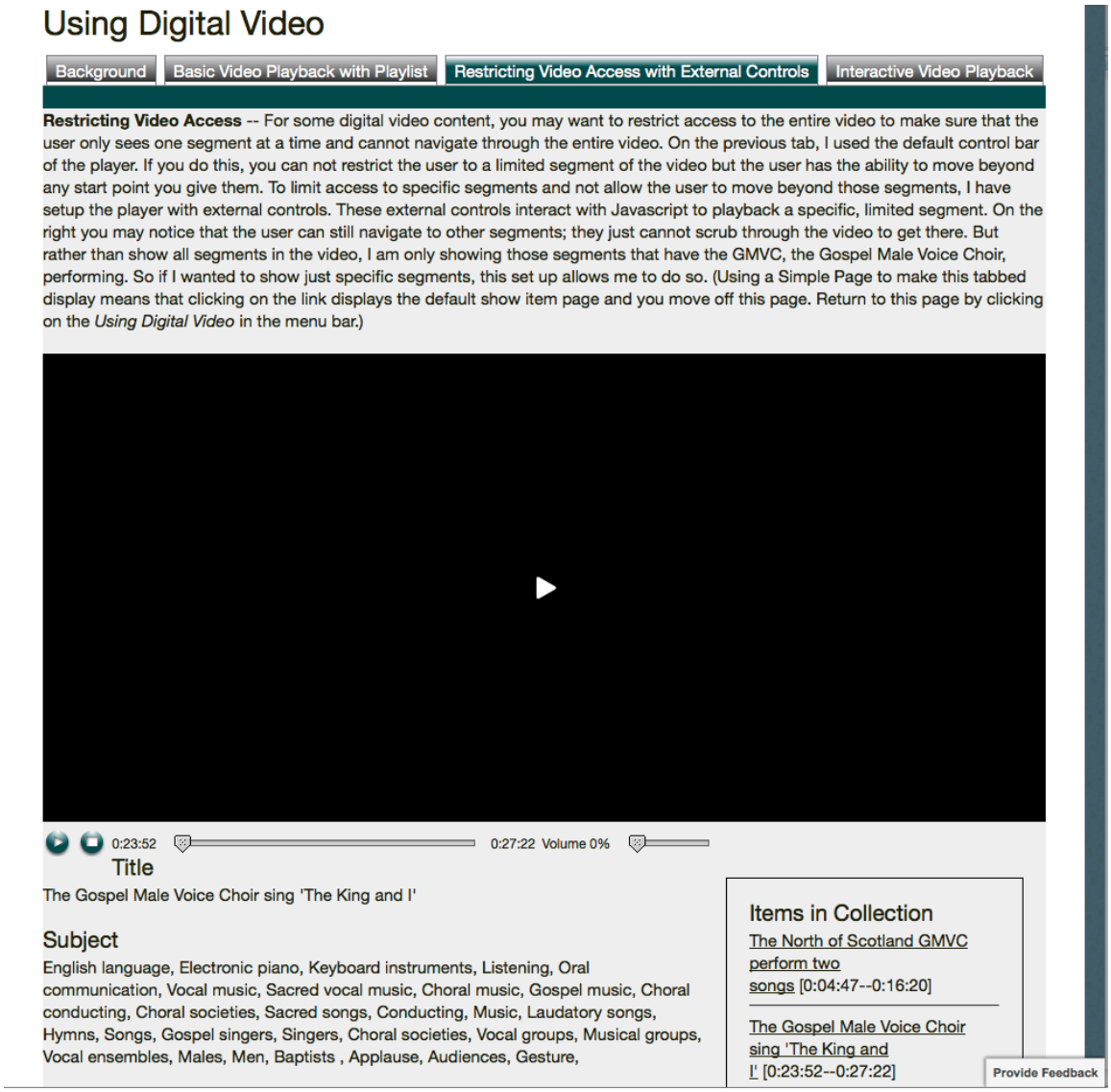
New Life sing 'Be Thou My Vision' [0:21:11--0:23:52]

The Gospel Male Voice Choir sing 'The King and I' [0:23:52--0:27:22]

Votes of thanks [0:29:52--0:32:03]

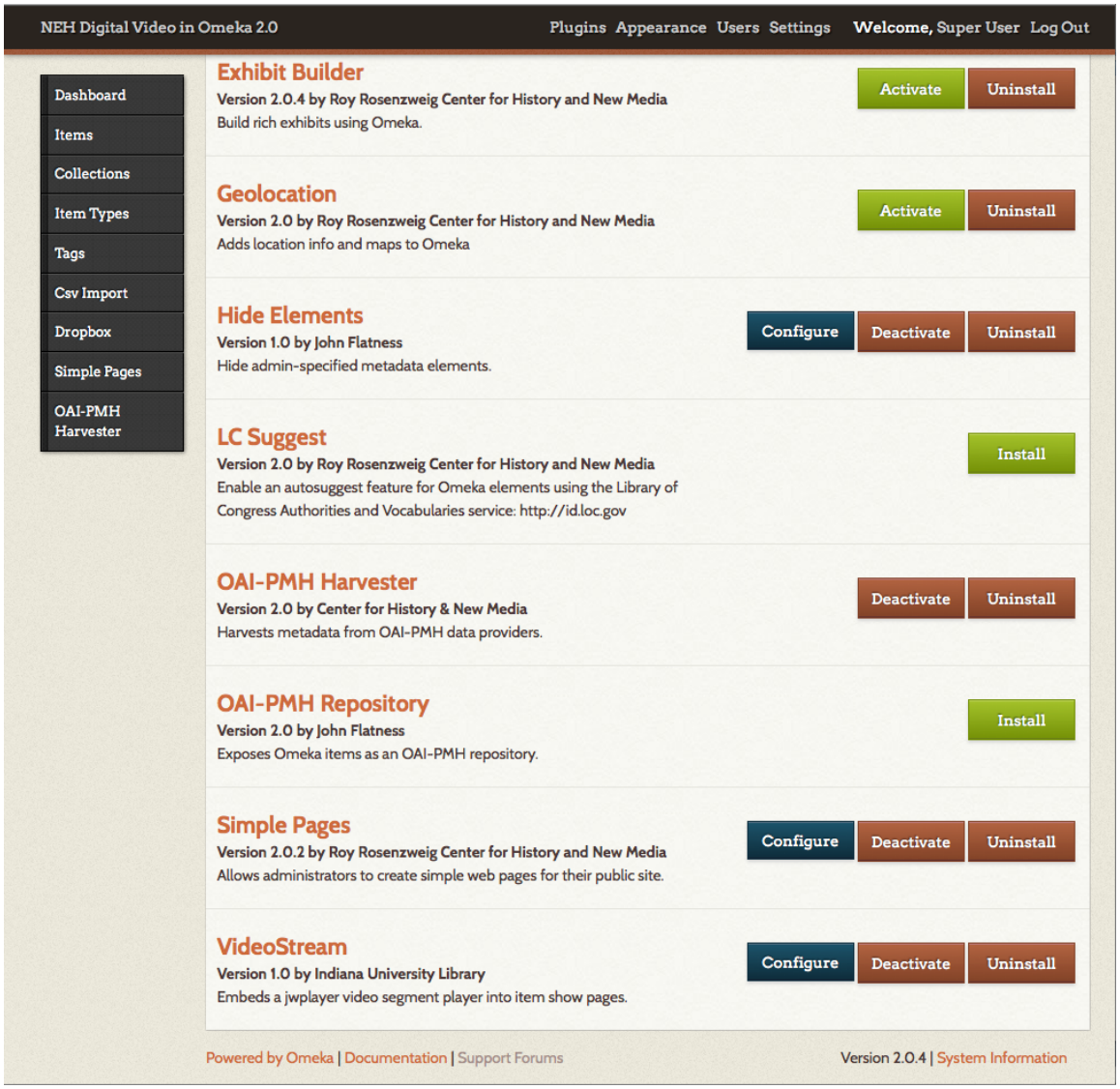
Congregation sing 'In the Sweet By-and By' [0:32:03--0:34:44]

This is a continuation of the previous page showing the Dublin Core metadata fields and down the right side a list of segments to select from.

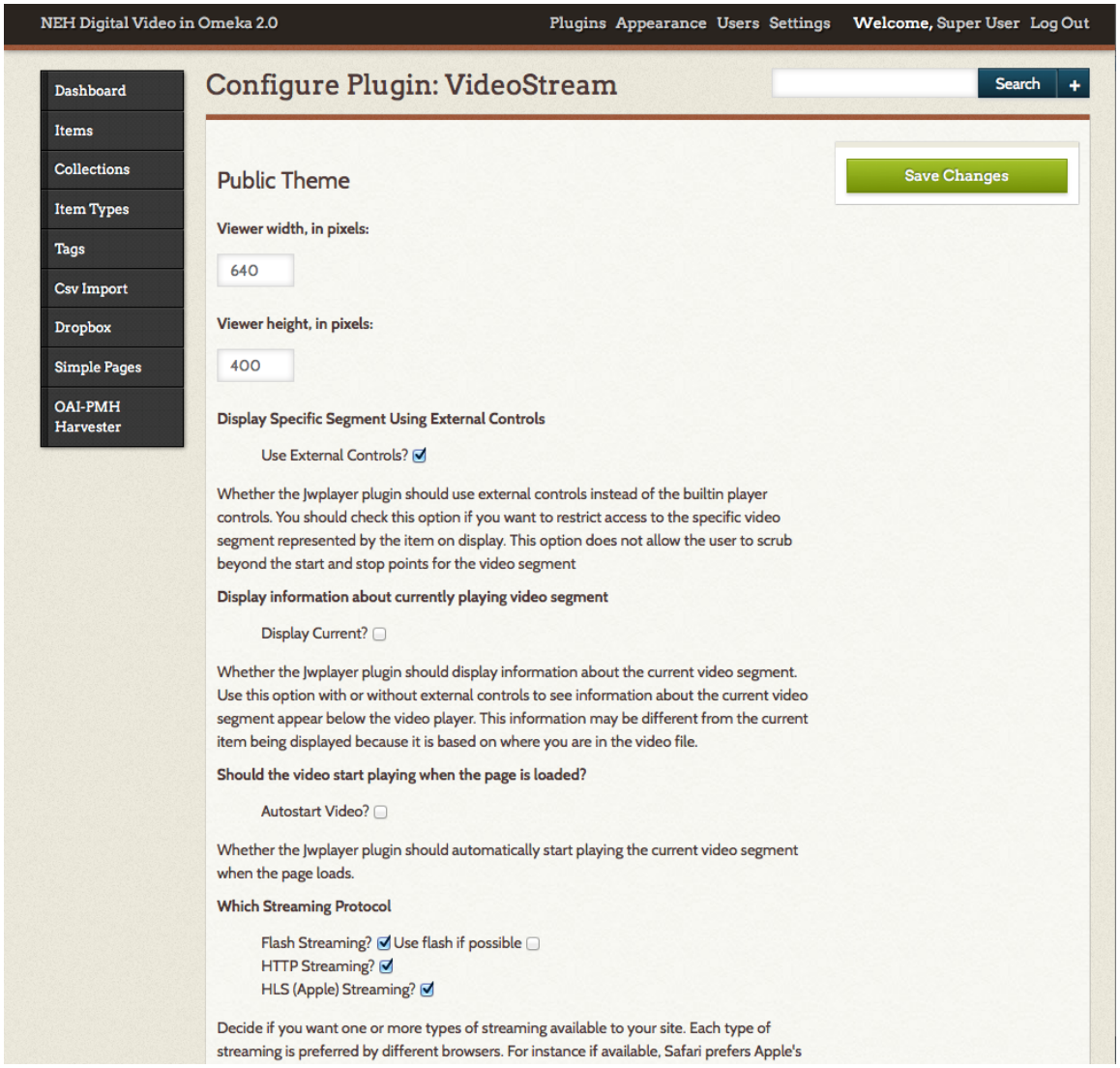


This theme supports individual video segment playback. The playback is restricted to a specific video segment and the user can not scrub before the start time or after the end time. Down the right side is a list to select other segments from.

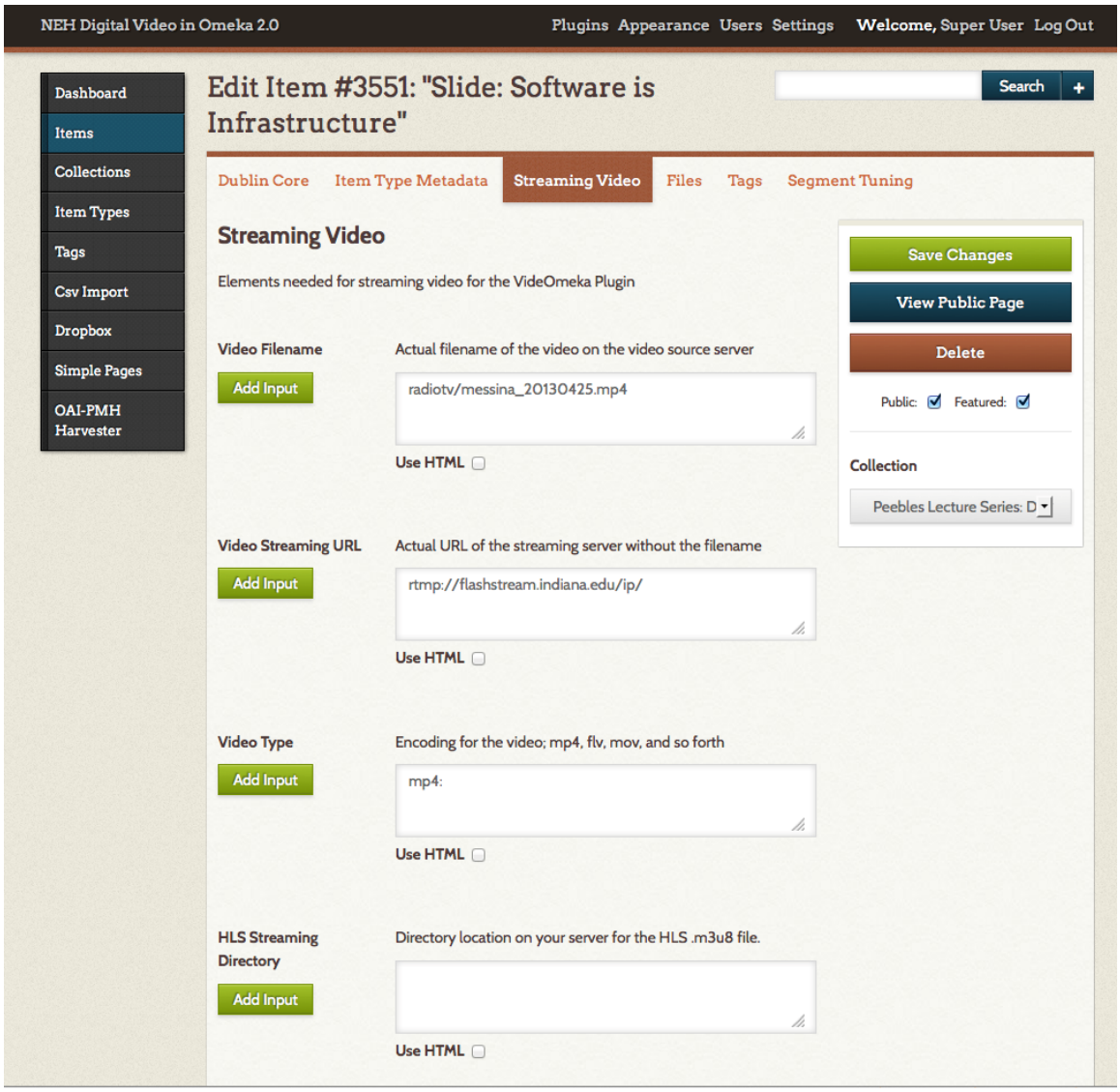
This theme support Simile Timeline to create an interactive approach to selecting a segment to play. Segments can be selected from the time line and playback is restricted to just that video segment. Information about the segment appears below the timeline.



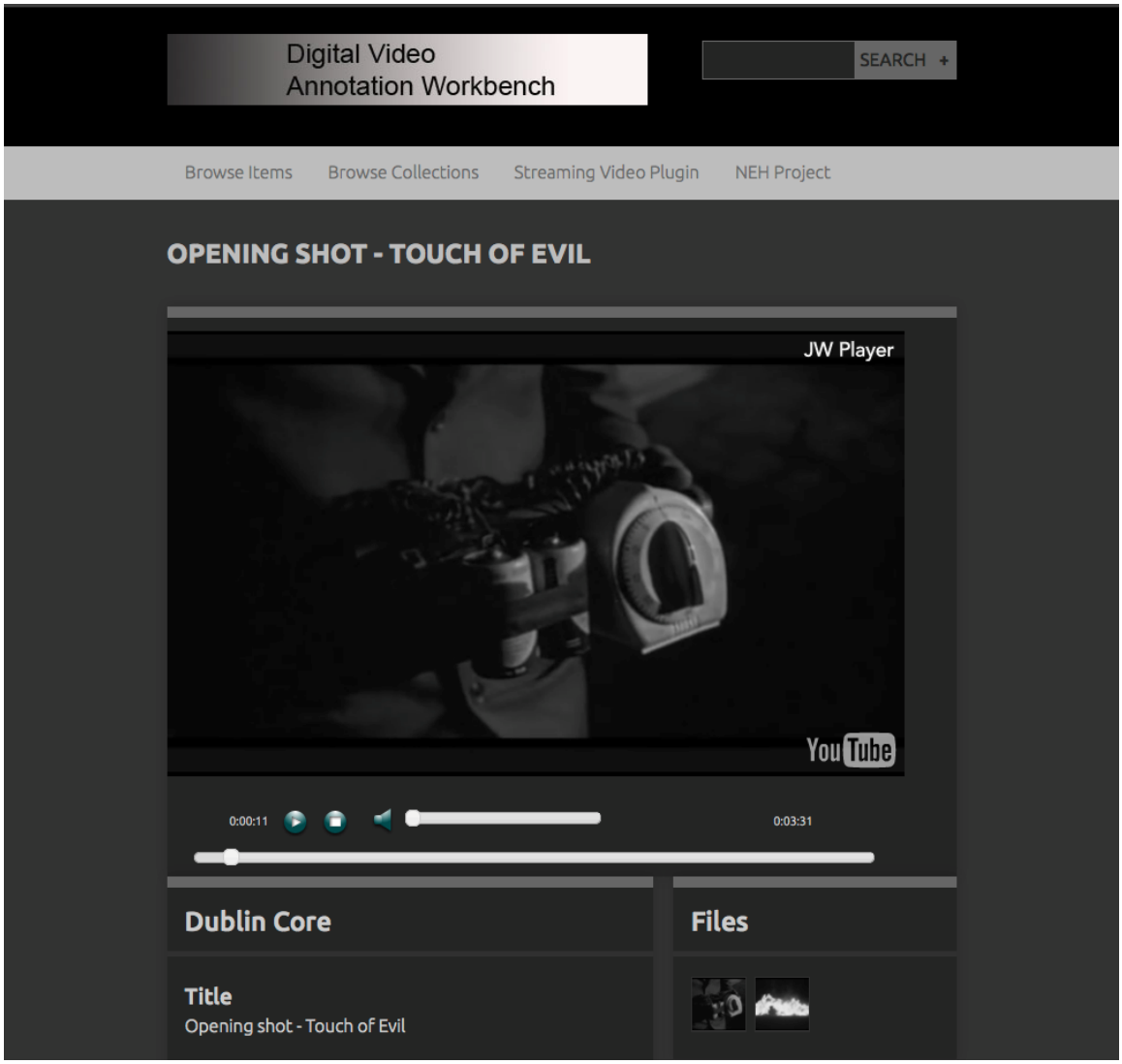
This is the plugin page for Omeak 2+. The VideoStream plugin is at the bottom of the page. When the user selects Configure, they would be taken the next image.



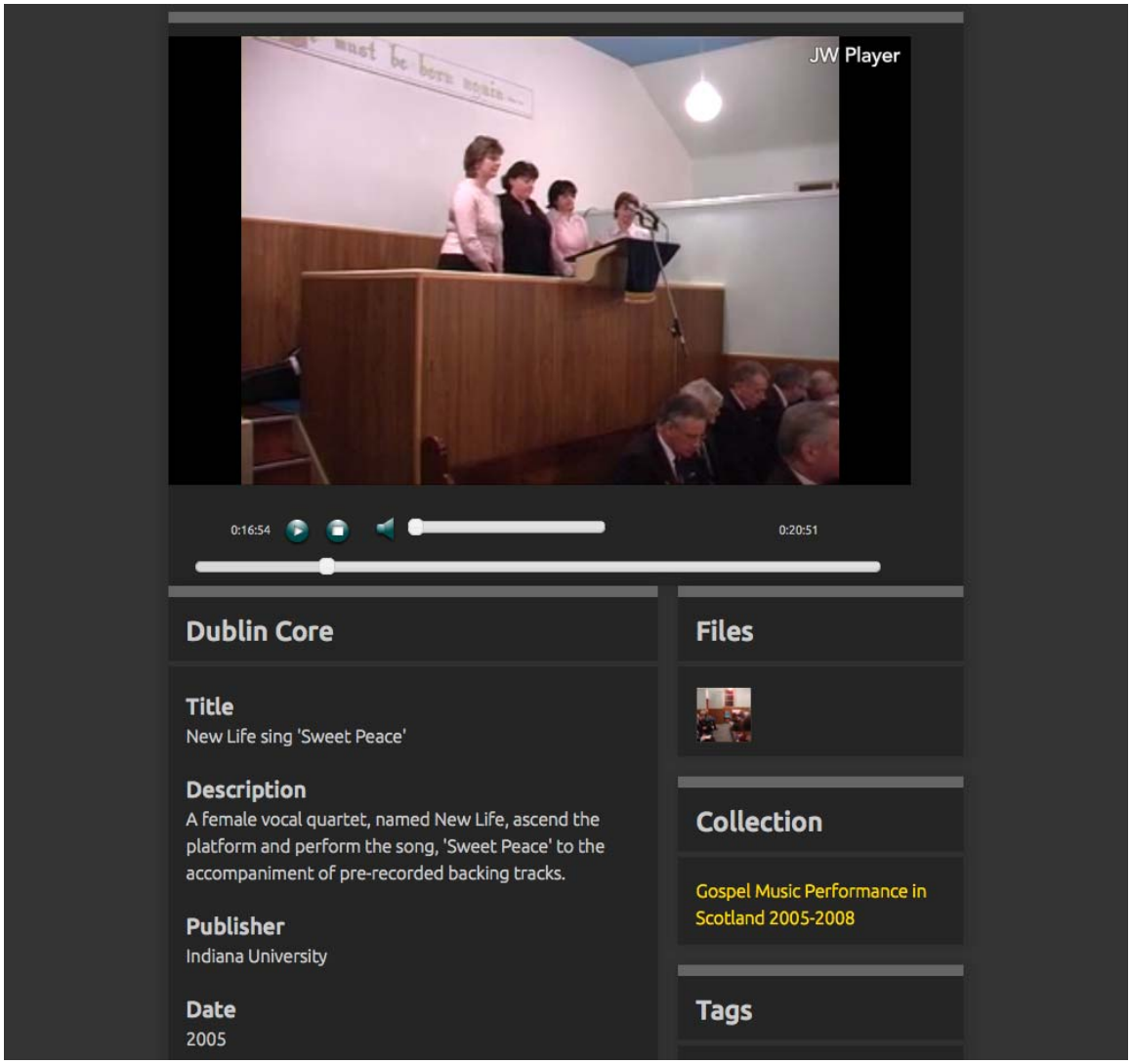
This page shows the major options the user can choose for the VideoStream plugin. There is documentation for this plugin at the http://www.dlib.indiana.edu/projects/omeka2 website.



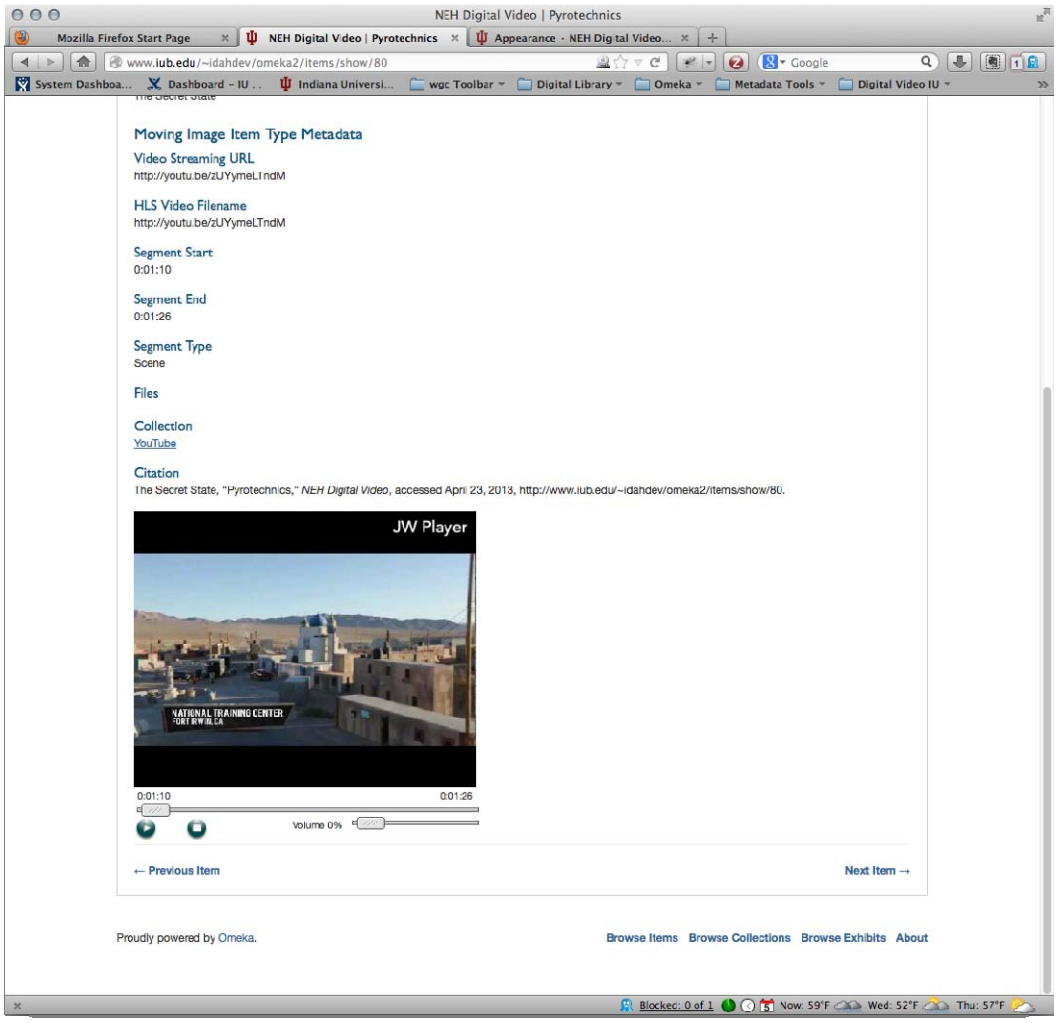
Once the VideoStream plugin is installed and configured, at the Edit Item page, the user will see a new tab for Streaming Video. This tab contains the Element Set created by the VideoStream plugin and contains fields for the video filename, location of the streaming video server, start time and end time of the segment and so forth.



This is the VideoStream plugin display on the show item page using the external controls, instead of the built-in controls using the Seasons theme.



Another view of the video playback on the show item page in Omeka using the Seasons theme.



This page is the video playback page using the Berlin theme.